Behavior, Purpose and Teleology

Arturo Rosenblueth, Norbert Wiener and Julian Bigelow

in: *Philosophy of Science*, **10(1943), S. 18–24.**

This essay has two goals. The first is to define the behavioristic study of natural events and to classify behavior. The second is to stress the importance of the concept of purpose.

Given any object, relatively abstracted from its surroundings for study, the behavioristic approach consists in the examination of the output of the object and of the relations of this output to the input. By output is meant any change produced in the surroundings by the object. By input, conversely, is meant any event external to the object that modifies this object in any manner.

The above statement of what is meant by the behavioristic method of study omits the specific structure and the instrinsic organization of the object. This omission is fundamental because on it is based the distinction between the behavioristic and the alternative functional method of study. In a functional analysis, as opposed to a behavioristic approach, the main goal is the intrinsic organization of the entity studied, its structure and its properties; the relations between the object and the surroundings are relatively incidental.

From this definition of the behavioristic method a broad definition of behavior ensues. By behavior is meant any change of an entity with respect to its surroundings. This change may be largely an output from the object, the input being then minimal, remote or irrelevant; or else the change may be immediately traceable to a certain input. Accordingly, any modification of an object, detectable externally, may be denoted as behavior. The term would be, therefore, too extensive for usefulness were it not that it may be restricted by apposite adjectives — i.e., that behavior may be classified.

The consideration of the changes of energy involved in behavior affords a basis for classification. Active behavior is that in which the object is the source of the output energy involved in a given specific reaction. The object may store energy supplied by a remote or relatively immediate input, but the input does not energize the output directly. In passive behavior, on the contrary, the object is not a source of energy; all the energy in the output can be traced to the immediate input (e.g., the throwlates the motor regions of the cerebral cortex he does not duplicate a voluntary reaction; he trips efferent, output pathways, but does not trip a purpose, as is done voluntarily.

The view has often been expressed that all machines are purposeful. This view is untenable. First may be mentioned mechanical devices such as a roulette, designed precisely for purposelessness. Then may be considered devices such as a clock, designed, it is true, with a purpose, but having a performance which, although orderly, is not purposeful — i.e., there is no specific final condition toward which the movement of the clock strives. Similarly, although a gun may be used for a definite purpose, the attainment of a goal is not intrinsic to the performance of the gun; random shooting can be made, deliberately purposeless.

Some machines, on the other hand, are intrinsically purposeful. A torpedo with a target-seeking mechanism is an example. The term servo-mechanisms has been coined precisely to designate machines with intrinsic purposeful behavior.

It is apparent from these considerations that although the definition of purposeful behavior is relatively vague, and hence operationally largely meaningless, the concept of purpose is useful and should, therefore, be retained.

Purposeful active behavior may be subdivided into two classes: *feed-back* (or *teleological*) and *non-feed-back* (or *non-teleological*) The expression feed-back is used by engineers in two different senses. In a broad sense it may denote that some of the output energy of an apparatus or machine is returned as input; an example is an electrical amplifier with feed-back. The feed-back is in these cases positive — the fraction of the output which reenters the object has the same sign as the original input signal. Positive feed-back adds to the input signals, it does not correct them. The term feed-back is also employed in a more restricted sense to signify that the behavior of an object is controlled by the margin of error at which the object stands at a given time with reference to a relatively specific goal. The feed-back is then negative, that is, the signals from the goal are used to restrict outputs which would otherwise go beyond the goal. It is this second meaning of the term feed-back that is used here.

All purposeful behavior may be considered to require negative feed-back. If a goal is to be attained, some signals from the goal are necessary at some time to direct the behavior. By non-feed-back behavior is meant that in which there are no signals from the goal which modify the activity of the object in the course of the behavior. Thus, a machine may be set to impinge upon a luminous object although the machine may be insensitive to light. Similarly, a snake may strike at a frog, or a frog at a fly, with no visual or other report from the prey after the movement has started. Indeed, the movement is in these cases so fast that it is not likely that nerve impulses would have time to arise at the retina, travel to the central nervous system and set up further impulses which would reach the muscles in time to modify the movement effectively.

As opposed to the examples considered, the behavior of some machines and some reactions of living organisms involve a continuous feed-back from the goal that modifies and guides the behaving object. This type of behavior is more effective than that mentioned above, particularly when the goal is not stationary. But continuous feedback control may lead to very clumsy behavior if the feed-back is inadequately damped and becomes therefore positive instead of negative for certain frequencies of oscillation. Suppose, for example, that a machine is designed with the purpose of impinging upon a moving luminous goal; the path followed by the machine is controlled by the direction and intensity of the light from the goal. Suppose further that the machine overshoots seriously when it follows a movement of the goal in a certain direction; an even stronger stimulus will then be delivered which will turn the machine in the opposite direction. If that movement again overshoots a series of increasingly larger oscillations will ensue and the machine will miss the goal.

This picture of the consequences of undamped feed-back is strikingly similar to that seen during the performance of a voluntary act by a cerebellar patient. At rest the subject exhibits no obvious motor disturbance. If he is asked to carry a glass of water from a table to his mouth, however, the hand carrying the glass will execute a series of oscillatory motions of increasing amplitude as the glass approaches his mouth, so that the water will spill and the purpose will not be fulfilled. This test is typical of the disorderly motor performance of patients with cerebellar disease. The analogy with the behavior of a machine with undamped feed-back is so vivid that we venture to suggest that the main function of the cerebellum is the control of the feed-back nervous mechanisms involved in purposeful motor activity.

Feed-back purposeful behavior may again

gling out of the class of predictive behavior, a class particularly interesting since it suggests the possibility of systematizing increasingly more complex tests of the behavior of organisms. It emphasizes the concepts of purpose and of teleology, concepts which, although rather discredited at present, are shown to be important. Finally, it reveals that a uniform behavioristic analysis is applicable to both machines and living organisms, regardless of the complexity of the behavior.

It has sometimes been stated that the designers of machines merely attempt to duplicate the performances of living organisms. This statement is uncritical. That the gross behavior of some machines should be similar to the reactions of organisms is not surprising. Animal behavior includes many varieties of all the possible modes of behavior and the machines devised so far have far from exhausted all those possible modes. There is, therefore a considerable overlap of the two realms of behavior. Examples, however, are readily found of man-made machines with behavior that transcends human behavior. A machine with an electrical output is an instance; for men, unlike the electric fishes, are incapable of emitting electricity. Radio transmission is perhaps an even better instance, for no animal is known with the ability to generate short waves, even if so-called experiments on telepathy are considered seriously.

A further comparison of living organisms and machines leads to the following inferences. The methods of study for the two groups are at present similar. Whether they should always be the same may depend on whether or not there are one or more qualitatively distinct, unique characteristics present in one group and absent in the other. Such qualitative differences have not appeared so far.

The broad classes of behavior are the same in machines and in living organisms. Specific, narrow classes may be found exclusively in one or the other. Thus, no machine is available yet that can write a Sanscrit-Mandarin dictionary. Thus, also, no living organism is known that rolls on wheels — imagine what the result would have been if engineers had insisted on copying living organisms and had therefore put legs and feet in their locomotives, instead of wheels.

While the behavioristic analysis of machines and living organisms is largely uniform, their functional study reveals deep differences. Structurally, organisms are mainly colloidal, and include prominently protein molecules, large, — complex and anisotropic; machines are chiefly metallic and include mainly simple molecules. From the standpoint of their energetics, machines usually exhibit relatively large differences of potential, which permit rapid mobilization of energy; in organisms the energy is more uniformly distributed, it is not very mobile. Thus, in electric machines conduction is mainly electronic, whereas in organisms electric changes are usually ionic.

Scope and flexibility are achieved in machines largely by temporal multiplication of effects; frequencies of one million per second or more are readily obtained and utilized. In organisms, spatial multiplication, rather than temporal, is the rule; the temporal achievements are poor — the fastest nerve fibers can only conduct about one thousand impulses per second; spatial multiplication is on the other hand abundant and admirable in its compactness. This difference is well illustrated by the comparison of a television receiver and the eye. The television receiver may be described as a single cone retina; the images are formed by scanning — i.e. by orderly successive detection of the signal with a rate of about 20 million per second. Scanning is a process which seldom or never occurs in organisms, since it requires fast frequencies for effective performance. The eye uses a spatial, rather than a temporal multiplier. Instead of the one cone of the television receiver a human eye has about 6.5 million cones and about 115 million rods.

If an engineer were to design a robot, roughly similar in behavior to an animal organism, he would not attempt at present to make it out of proteins and other colloids. He would probably build it out of metallic parts, some dielectrics and many vacuum tubes. The movements of the robot could readily be much faster and more powerful than those of the original organism. Learning and memory, however, would be quite rudimentary. In future years, as the knowledge of colloids and proteins increases, future engineers may attempt the design of robots not only with a behavior, but also with a structure similar to that of a mammal. The ultimate model of a cat is of course another cat, whether it be born of still another cat or synthesized in a laboratory.

In classifying behavior the term *teleology* was used as synonymous with *purpose controlled by feed-back*. Teleology has been interpreted in the past to imply purpose and the vague concept of a *final cause* has been often added. This concept of final causes has led to the opposition of teleology to determinism. A discussion of causality, determinism and final causes is beyond the scope of this essay. It may be pointed out, however, that purposefulness, as defined here, is quite independent of causality, initial or final. Teleology has been discredited chiefly because it was defined to imply a cause subsequent in time to a given effect. When this aspect of teleology was dismissed, however, the associated recognition of the importance of purpose was also unfortunately discarded. Since we consider purposefulness a concept necessary for the understanding of certain modes of behavior we suggest that a teleological study is useful if it avoids problems of causality and concerns itself merely with an investigation of purpose.

We have restricted the connotation of teleological behavior by applying this designation only to purposeful reactions which are controlled by the error of the reaction — i.e., by the difference between the state of the behaving object at any time and the final state interpreted as the purpose. Teleological behavior thus becomes synonymous with behavior controlled by negative feedback, and gains therefore in precision by a sufficiently restricted connotation.

According to this limited definition, teleology is not opposed to determinism, but to non-teleology. Both teleological and non-teleological systems are deterministic when the behavior considered belongs to the realm where determinism applies. The concept of teleology shares only one thing with the concept of causality: a time axis. But causality implies a one-way, relatively irreversible functional relationship, whereas teleology is concerned with behavior, not with functional relationships.